A POSTERIORI ERROR ESTIMATES FOR PARABOLIC PROBLEMS VIA ELLIPTIC RECONSTRUCTION AND DUALITY

OMAR LAKKIS AND CHARALAMBOS MAKRIDAKIS

Abstract. We use the elliptic reconstruction technique in combination with a duality approach to prove a posteriori error estimates for fully discrete backward Euler scheme for linear parabolic equations. As an application, we combine our result with the residual based estimators from the a posteriori estimation for elliptic problems to derive space-error estimators and thus a fully practical version of the estimators bounding the error in the $\mathrm{L}_\infty(0,T;\mathrm{L}_2(\Omega))$ norm. These estimators, which are of optimal order, extend those introduced by Eriksson and Johnson [EJ91] by taking into account the error induced by the mesh changes and allowing for a more flexible use of the elliptic estimators. For comparison with previous results, an application of our abstract results using residual estimators is provided.

1. Introduction

A posteriori error estimators and their use to derive adaptive mesh refinement algorithms to solve time-dependent problems constitute the object of intense research. The problem is appealing for the theorician as a test ground for novel analytical techniques as well as for the practitioners which are interested in minimizing the ammount of computational time in order to obtain a satisfactory accuracy in the computer simulations of time-dependent PDE's. Both the theorical and practical aspects of a posteriori-based adaptive numerical methods for evolution partial differential equations has benefitted immensely from the surge in the production of dedicated papers in the last 15 years.

In a recent article [LM06] we have introduced the *elliptic reconstruction* for fully discrete explicit Euler schemes and we have applied it using the energy techniques to derive optimal-order a posteriori residual-based error estimators. In particular we were able to establish estimates for the error in the $\mathrm{L}_\infty(0,T;\mathrm{L}_2(\Omega))$ without the use of duality methods. Previous work for the spatially-discrete and time-continuous scheme was introduced by Makridakis & Nochetto [MN03].

Our chief goal in this paper is to explore the possiblity of using the *elliptic reconstruction* techinque in conjunction with the *duality* technique as introduced by Eriksson & Johnson [EJ91]. The *duality* technique provides an important alternative to energy techniques and is widely used for the derivation of a priori and a posteriori error estimates both for elliptic and parabolic problems. Since being first considered by Eriksson and Johnson in 1991 [EJ91] it has been developed in many different directions, including its use in *implicit and goal oriented a posteriori error estimates*.

Date: February 1, 2008.

2000 *Mathematics Subject Classification.* Primary: 65N30.

This work was partially supported by the E.U. RTN *Hyke* HPRN-CT-2002-00282 and the Marie Curie Fellowship Foundation.

The elliptic reconstruction has been used in combination with energy estimates, in which one mimicks the energy estimates for the parabolic equation in order to derive error estimates from a PDE where the error, or part thereof, is the "unknown". Our chief objective in this paper is to exhibit the flexibility of the elliptic reconstruction technique by showing that it can be completely decoupled from energy considerations. This is not obvious, indeed, in many works a posteriori analysis, the elliptic part is entangled with the parabolic part and there is not a clear cut difference between elliptic and parabolic effects. As noted in recent work on a posteriori analysis for time-dependent problems [AMN05, BBM05, BV04, e.g.], understanding the splitting between the elliptic, stationary, and parabolic, time-dependent, errors, as well as the part of the error where these effects are coupled, is important in designing adaptive methods.

An important by-product of our approach is that the mesh-change in time is considered as part of the proofs of our theorems. Indeed, unlike former derivations a posteriori error estimates via duality [EJ91], we do not impose on the mesh any assumption that are susceptible of violation in a practical implementation of the scheme, such as the no-refinement assumptions. We stress that once the elliptic reconstruction paraphernalia is introduced the analysis becomes quite clear and straightforward and the heart of the matter does not exceed more than 3 pages, in contrast with the more involved approach of Eriksson & Johnson.

To justify completely our effort from a practical side, we give an application of our theory with a concrete example where residual-based estimators are used [AO00, LN03]. We emphasize, however, that our results are not limited to the use of residual-based estimators and that other estimators which work for the L_2 norms in elliptic problems could be used.

1.1. **Basic set-up.** We start by introducing the exact PDE whose discretization is the object of this paper. Let Ω be a bounded domain of the Euclidean space $\mathbb{R}^d$, $d \in \mathbb{Z}^+$ and $T \in \mathbb{R}^+$. We assume throughout the paper that Ω is a polygonal convex domain, noticing that all the results can be extended to certain non-convex domains, like domains with reentrant corners in $d = 2$.

Given a Lebesgue measurable set $D \subset \mathbb{R}^d$, we use the following notation

$$\langle \phi, \psi \rangle_D := \int_D \phi(\boldsymbol{x})\psi(\boldsymbol{x})\,\mathrm{d}\mu(\boldsymbol{x}), \tag{1.1}$$

$$\|\phi\|_D := \|\phi\|_{\mathrm{L}_2(D)} := \langle \phi, \phi \rangle_D^{1/2}, \tag{1.2}$$

$$|\phi|_{k,D} := \left\| \mathrm{D}^k \phi \right\|_D, \text{ for } k \in \mathbb{Z}^+ \quad (\text{with } \mathrm{D}^1\phi := \nabla\phi, \text{ etc.}), \tag{1.3}$$

$$\|\phi\|_{k,D} := \left(\|\phi\|_D^2 + \sum_{j=1}^{k} |\phi|_{j,D}^2 \right)^{1/2}, \text{ for } k \in \mathbb{Z}^+, \tag{1.4}$$

where $\mathrm{d}\mu(\boldsymbol{x})$ is either the Lebesgue measure element $\mathrm{d}\boldsymbol{x}$, if D is has positive such measure, or the $(d-1)$-dimensional (Hausdorff) measure $\mathrm{d}s(\boldsymbol{x})$, if D has zero Lebesgue measure. In many instances, in order to compress notation and when there is no danger of engendering confusion, we may drop altogether the "differential" symbol from integrals. This convention applies also to integrals in time.

We will use the standard function spaces $\mathrm{L}_2(D)$, $\mathrm{H}^k(D)$, $\mathrm{H}^k_0(D)$ and denote by $\mathrm{H}^{-1}(D)$ the dual space of $\mathrm{H}^1_0(D)$ with the corresponding pairing written as $\langle \cdot \mid \cdot \rangle_D$. We omit the subscript D whenever $D = \Omega$. We denote the Poincaré constant relative to Ω by $C_{2,1}$ and take $|\cdot|_1$ to be the norm of $\mathrm{H}^1_0(\Omega)$. We use the usual duality identification

$$\mathrm{H}^1_0(\Omega) \subset \mathrm{L}_2(\Omega) \sim \mathrm{L}_2(\Omega)' \subset \mathrm{H}^{-1}(\Omega) \tag{1.5}$$

and the dual norm

$$\|\psi\|_{-1} := \sup_{0\neq\phi\in \mathrm{H}_0^1(\Omega)} \frac{\langle \psi \mid \phi\rangle}{|\phi|_1} \left(= \sup_{0\neq\phi\in \mathrm{H}_0^1(\Omega)} \frac{\langle \psi, \phi\rangle}{|\phi|_1}, \text{ if } \psi \in \mathrm{L}_2(\Omega)\right). \tag{1.6}$$

Let a be the elliptic bilinear form defined on $\mathrm{H}_0^1(\Omega)$ by

$$a(v,\psi) := \langle \boldsymbol{A}\nabla v, \nabla\psi\rangle, \ \forall v,\psi \in \mathrm{H}_0^1(\Omega) \tag{1.7}$$

where "∇" denotes the spatial gradient and the matrix $\boldsymbol{A} \in \mathrm{L}_\infty(\Omega)^{d\times d}$ is such that

$$a(\psi,\phi) \le \beta\, |\psi|_1\, |\phi|_1, \quad \forall \phi,\psi \in \mathrm{H}_0^1(\Omega), \tag{1.8}$$

$$a(\phi,\phi) \ge \alpha\, |\phi|_1^2, \quad \forall \phi \in \mathrm{H}_0^1(\Omega), \tag{1.9}$$

with $\alpha,\beta \in \mathbb{R}^+$. We also use the *energy norm* $|\cdot|_a$ defined as

$$|\phi|_a := a(\phi,\phi)^{1/2}, \ \forall\phi \in \mathrm{H}_0^1(\Omega). \tag{1.10}$$

It is equivalent to the norm $|\cdot|_1$ on the space $\mathrm{H}_0^1(\Omega)$, in view of (1.8) and (1.9). In particular, we will often use the following inequality

$$|\phi|_1 \le \alpha^{-1/2}\, |\phi|_a, \ \forall\phi \in \mathrm{H}_0^1(\Omega). \tag{1.11}$$

Let $u \in \mathrm{L}_\infty(0,T;\mathrm{H}_0^1(\Omega))$, with $\partial_t u \in \mathrm{L}_2(0,T;\mathrm{H}^{-1}(\Omega))$, be the unique solution of the linear parabolic problem

$$\begin{aligned} &\langle \partial_t u \mid \phi\rangle + a(u,\phi) = \langle f,\phi\rangle \quad \forall\phi \in \mathrm{H}_0^1(\Omega),\\ &\text{and } u(0) = g, \end{aligned} \tag{1.12}$$

where $f \in \mathrm{L}_2(\Omega\times(0,T))$ and $g \in \mathrm{H}_0^1(\Omega)$.

Whenever not stated explicitly, we assume that the data $f, g, \boldsymbol{A}$ and the solution u of the above problem are sufficiently regular for our purposes.

In order to discretize the time variable in (1.12), we introduce the partition $0 = t_0 < t_1 < \ldots < t_N = T$ of $[0,T]$. Let $I_n := (t_{n-1}, t_n]$ and we denote by $\tau_n := t_n - t_{n-1}$ the time steps. We will consistenly use the following "superscript convention": whenever a function depends on time, e.g. $f(\boldsymbol{x},t)$, and the time is fixed to be $t = t_n$, $n \in [0:N]$ we denote it by $f^n(\boldsymbol{x})$. Moreover, we often drop the space dependence explictly, e.g, we write $f(t)$ and f^n in reference to the previous sentence.

We use a standard FEM to discretize the space variable. Let $(\mathscr{T}_n)_{n\in[0:N]}$ be a family of conforming triangulations of the domain Ω [BS94, Cia78]. These triangulations are allowed to change arbitrarily from a timestep to the next, as long as they maintain some very mild *compatibility* requirements. Our use of the term "compatibility" is precisely defined in [LM06]; it is an extremely mild requirement which is easily implemented in practice.

For each given a triangulation $\mathscr{T}_n$, we denote by h_n its meshsize function defined as

$$h_n(\boldsymbol{x}) = \operatorname{diam}(K), \text{ where } K \in \mathscr{T}_n \text{ and } \boldsymbol{x} \in K, \tag{1.13}$$

for all $\boldsymbol{x} \in \Omega$. We also denote by $\mathscr{S}_n$ the set of *internal sides* of $\mathscr{T}_n$, these are edges in $d = 2$—or faces in $d = 3$—that are contained in the interior of Ω; the *interior mesh of edges* Σ_n is then defined as the union of all internal sides $\cup_{E\in\mathscr{S}_n} E$. We associate with these triangulations the *finite element spaces*:

$$\mathbb{V}^n := \left\{\phi \in \mathrm{H}_0^1(\Omega) : \forall K \in \mathscr{T}_n : \phi|_K \in \mathbb{P}^\ell\right\} \tag{1.14}$$

where $\mathbb{P}^\ell$ is the space of polynomials in d variables of degree at most $\ell \in \mathbb{Z}^+$. Given two successive compatible triangulations $\mathscr{T}_{n-1}$ and $\mathscr{T}_n$, we define $\hat{h}_n := \max(h_n, h_{n-1})$ [LM06, Appendix]. We will also use the sets $\hat{\Sigma}_n := \Sigma_n \cap \Sigma_{n-1}$ and $\check{\Sigma}_n := \Sigma_n \cup \Sigma_{n-1}$.

1.2. **Definition** (fully discrete scheme) We consider the following fully discrete scheme of problem (1.12) associated with the finite element spaces $\mathbb{V}^n$:
(1.15)

$$U^0 := I^0 u(0), \text{ and}$$
$$\tau_n^{-1} \left\langle U^n - U^{n-1}, \Phi_n \right\rangle + a\left(U^n, \Phi_n\right) = \left\langle \tilde{f}^n, \Phi_n \right\rangle, \quad \forall \Phi_n \in \mathbb{V}^n, \text{for } n \in [1:N].$$

Here the operator I^0 is some suitable interpolation or projection operator from $\mathrm{H}_0^1(\Omega)$, or $\mathrm{L}_2(\Omega)$, onto $\mathbb{V}^n$, and $\tilde{f}^n$ equals either the value of f at t_n, $f^n := f(\cdot, t_n)$, or its time-average on I_n, $\int_{t_{n-1}}^{t_n} f(\cdot, t)\,\mathrm{d}t/(t_n - t_{n-1})$. This schme is the standard backward (or implicit) Euler–Galerkin finite element scheme [Tho97].

In the sequel we shall use a continuous piecewise linear extension in time of the sequence (U^n) which we denote by $U(t)$ for $t \in [0,T]$ (see §2.4 for the precise definition).

1.3. **A posteriori estimates and reconstruction operators.** Suppose we associate with U an auxiliary function $\omega : [0,T] \to \mathrm{H}_0^1(\Omega)$, in such a way that the *total error*

$$e := U - u \tag{1.16}$$

can be decomposed as follows

$$e = \rho - \epsilon \tag{1.17}$$

$$\epsilon := \omega - U, \quad \rho := \omega - u. \tag{1.18}$$

The new auxiliary function ω is *reconstructed* from the given approximation U. The success of this splitting for the estimates to follow rests in the following properties:

1. The error ϵ is easily controlled by a posteriori quantities of optimal order.
2. The error ρ satisfies a modification of the original PDE whose right-hand side depends on ϵ and U. This right-hand side can be bounded a posteriori in an optimal way.

Therefore in order to successfully apply this idea we must select a suitable reconstructed function ω. In our case, this choice is dictated by the elliptic operator at hand; the precise definition is given in §2.2. In addition the effect of mesh modification will reflect in the right-hand side of the equation for ρ. As a result of our choice for ω we are able to derive optimal order estimators for the error in $\mathrm{L}_\infty(0,T;\mathrm{L}_2(\Omega))$, as well as in $\mathrm{L}_\infty(0,T;\mathrm{H}_0^1(\Omega))$ and $\mathrm{H}^1(0,T;\mathrm{L}_2(\Omega))$. In addition, our choosing ω as the elliptic reconstruction will have the effect of separating the spatial approximation error from the time approximation as much as possible. We show that the spatial approximation is embodied in ϵ which will be referred to as the *elliptic reconstruction error* whereas the time approximation error information is conveyed by ρ, a fact that motivates the name *main parabolic error* for this term. This "splitting" of the error is already apparent in the spatially discrete case [MN03].

With the above notation, we prove in the sequel that ρ satisfies the following variational equation.

1.4. **Lemma** (main parabolic error equation) *For each $n \in [1:N]$, and for each $\phi \in \mathrm{H}_0^1(\Omega)$,*

$$\begin{aligned} \langle \partial_t \rho \,|\, \phi \rangle + a\left(\rho, \phi\right) = \langle \partial_t \epsilon, \phi \rangle + a\left(\omega - \omega^n, \phi\right) \\ + \langle P_0^n f^n - f, \phi \rangle + \tau_n^{-1} \left\langle P_0^n U^{n-1} - U^{n-1}, \phi \right\rangle \text{ on } I_n. \end{aligned} \tag{1.19}$$

Here P_0^n denotes the L^2-projection into $\mathbb{V}^n$.

1.5. **Outline.** The rest of this article is organized as follows: in §2 we introduce the main tools related to the elliptic resconstruction, in §3 we outline the strategy of the proof in section §4 we present an application using the residual-based estimators.

2. The discrete scheme and the elliptic reconstruction

We introduce the numerical schemes that we study, some basic tools including the definition of the elliptic reconstruction. To keep this section brief, and due to a substantial overlap, we refer the reader to our previous paper for the details [].

Although the definitions some in parts of this section are independent of the time discretization and could be applied to any finite element space, we still use the space $\mathbb{V}^n$ defined in the introduction.

2.1. **Definition** (representation of the elliptic operator, discrete elliptic operator, projections) Suppose a function $v \in \mathbb{V}^n$, the bilinear form can be then represented as

$$a(v, \phi) = \sum_{K \in \mathscr{T}_n} \langle -\operatorname{div}(\boldsymbol{A}\nabla v), \phi\rangle_K + \sum_{E \in \mathscr{S}_n} \langle J[v], \phi\rangle_E, \ \forall \phi \in \mathrm{H}^1_0(\Omega), \tag{2.1}$$

where $J[v]$ is the *spatial jump of the field $\boldsymbol{A}\nabla v$ across an element side $E \in \mathscr{S}_n$* defined as

$$J[v]|_E(\boldsymbol{x}) = [\![\boldsymbol{A}\nabla v]\!]_E(\boldsymbol{x}) := \lim_{\varepsilon \to 0} (\boldsymbol{A}\nabla v(\boldsymbol{x} + \varepsilon\boldsymbol{\nu}_E) - \boldsymbol{A}\nabla v(\boldsymbol{x} - \varepsilon\boldsymbol{\nu}_E)) \cdot \boldsymbol{\nu}_E \tag{2.2}$$

where $\boldsymbol{\nu}_E$ is a choice, which does not influence this definition, between the two possible normal vectors to E at the point $\boldsymbol{x}$.

Since we use the representation (2.1) quite often, we introduce now a practical notation that makes it shorter and thus easier to manipulate in convoluted computations. For a finite element function, $v \in \mathbb{V}^n$ (or more generally for any Lipschitz continuous function v that is $\mathrm{C}^2(\operatorname{int}(K))$, for each $K \in \mathscr{T}_n$), denote by $A_{\mathrm{el}}v$ the *regular part* of the distribution $-\operatorname{div}(\boldsymbol{A}\nabla v)$, which is defined as a piecewise continuous function such that

$$\langle A_{\mathrm{el}}v, \phi\rangle = \sum_{K \in \mathscr{T}_n} \langle -\operatorname{div}(\boldsymbol{A}\nabla v), \phi\rangle, \quad \forall \phi \in \mathrm{H}^1_0(\Omega). \tag{2.3}$$

The operator A_{el} is sometime refered to as the *elementwise elliptic operator*, as it is the result of the application of $-\operatorname{div}(\boldsymbol{A}\nabla\cdot)$ only on the interior of each element $K \in \mathscr{T}_n$. This observation justifies our subscript in the notation. We shall write the representation (2.1) in the shorter form

$$a(v, \phi) = \langle A_{\mathrm{el}}v, \phi\rangle + \langle J[v], \phi\rangle_{\Sigma_n}, \quad \forall \phi \in \mathrm{H}^1_0(\Omega). \tag{2.4}$$

Let us now recall some more basic definitions that we will be using. The *discrete elliptic operator* associated with the bilinear form a and the finite element space $\mathbb{V}^n$ is the operator $A^n : \mathrm{H}^1_0(\Omega) \to \mathbb{V}^n$ defined by

$$\langle A^n v, \phi_n\rangle = a(v, \phi_n), \quad \forall \phi_n \in \mathbb{V}^n, \tag{2.5}$$

for $v \in \mathrm{H}^1_0(\Omega)$. The L_2-*projection operator* is defined as the operator $P^n_0 : \mathrm{L}_2(\Omega) \to \mathbb{V}^n$ such that

$$\langle P^n_0 v, \phi_n\rangle = \langle v, \phi_n\rangle, \quad \forall \phi_n \in \mathbb{V}^n \tag{2.6}$$

for $v \in \mathrm{L}_2(\Omega)$; and the *elliptic projection operator* $P^n_1 : \mathrm{H}^1_0(\Omega) \to \mathbb{V}^n$ is defined by

$$a(P^n_1 v, \phi_n) = a(v, \phi_n), \quad \forall \phi_n \in \mathbb{V}^n. \tag{2.7}$$

The elliptic reconstruction, which we define next, is a partial right inverse of the elliptic projection [MN03]. (Notice that a similar operator has been introduced by García-Archilla & Titi, albeit with different applications in mind than ours [GAT00].)

2.2. **Definition** (elliptic reconstruction) We define the *elliptic reconstruction operator* associated with the bilinear form a and the finite element space $\mathbb{V}^n$ to be the unique operator $\mathscr{R}^n : \mathrm{H}_0^1(\Omega) \to \mathrm{H}_0^1(\Omega)$ such that

$$a\left(\mathscr{R}^n v, \phi\right) = \left\langle A^n v, \phi\right\rangle, \quad \forall \phi \in \mathrm{H}_0^1(\Omega), \tag{2.8}$$

for a given $v \in \mathrm{H}_0^1(\Omega)$. The function $\mathscr{R}^n v$ is referred to as the *elliptic reconstruction* of v. A crucial property of $\mathscr{R}^n$ is that $v - \mathscr{R}^n v$ is orthogonal, with respect to a, to $\mathbb{V}^n$:

$$a\left(v - \mathscr{R}^n v, \phi_n\right) = 0, \quad \forall \phi_n \in \mathbb{V}^n. \tag{2.9}$$

From this property and by now standard techniques in a posteriori error estimates for elliptic problems [AO00, Bra01, Ver96], it is possible to obtain the following result.

2.3. **Lemma** (elliptic reconstruction error estimates) *For any $v \in \mathbb{V}^n$ the following estimates hold true*

$$\left|\mathscr{R}^n v - v\right|_1 \le \frac{C_{3,1}}{\alpha}\left\|(A_{\mathrm{el}} v - A^n v) h_n\right\| + \frac{C_{5,1}}{\alpha}\left\|J[v] h_n^{1/2}\right\|_{\Sigma_n}, \tag{2.10}$$

$$\left\|\mathscr{R}^n v - v\right\| \le C_{6,2}\left\|(A_{\mathrm{el}} v - A^n v) h_n^2\right\| + C_{10,2}\left\|J[v] h_n^{3/2}\right\|_{\Sigma_n}, \tag{2.11}$$

where the constants $C_{k,j}$ are defined in [LM06, Appendix B].

We will use the following shorthand

$$\omega = \mathscr{R} U. \tag{2.12}$$

2.4. **Definition** (discrete time extensions and derivatives) Given any discrete function of time—that is, a sequence of values associated with each time node t_n—e.g., (U^n), we associate to it the continuous function of time defined by the Lipschitz continuous piecewise linear interpolation, e.g.,

$$U(t) := l_{n-1}(t) U^{n-1} + l_n(t) U^n, \text{ for } t \in I_n \text{ and } n \in [1:N]; \tag{2.13}$$

where the functions l are the hat functions defined by

$$l_n(t) := \frac{t - t_{n-1}}{\tau_n} \mathbf{1}_{I_n}(t) - \frac{t - t_{n+1}}{\tau_{n+1}} \mathbf{1}_{I_{n+1}}(t), \text{ for } t \in [0,T] \text{ and } n \in [0:N], \tag{2.14}$$

$\mathbf{1}_X$ denoting the characteristic function of the set X. The *time-dependent elliptic reconstruction* of U is the function

$$\omega(t) := l_{n-1}(t) \mathscr{R}^{n-1} U^{n-1} + l_n(t) \mathscr{R}^n U^n, \text{ for } t \in I_n \text{ and } n \in [1:N]. \tag{2.15}$$

Notice that ω is a Lipschitz continuous function of time.

We introduce the following definitions whose purpose is to make notation more compact:

(a) *Discrete (backward) time derivative*

$$\partial U^n := \frac{U^n - U^{n-1}}{\tau_n} \tag{2.16}$$

Notice that $\partial U^n = \partial_t U(t)$, for all $t \in I_n$, hence we can think of ∂U^n as being the value of a discrete function at t_n. We thus define ∂U as the piecewise linear extension of $(\partial U^n)_n$, as we did with U.

(b) *Discrete (centered) second time derivative*

$$\partial^2 U^n := \frac{\partial U^{n+1} - \partial U^n}{\tau_n}. \tag{2.17}$$

(c) *Averaged (L_2-projected) discrete time derivative*

$$\overline{\partial} U^n := P_0^n \partial U^n = \frac{U^n - P_0^n U^{n-1}}{\tau_n}, \quad \forall n \in [1:N]\,. \tag{2.18}$$

The reason we introduce this notation for is that ∂U^n, in general, does not belong to the current finite element space, $\mathbb{V}^n$, whereas $\overline{\partial} U^n$ does.

2.5. **Remark** (pointwise form) The discrete elliptic operators A^n can be employed to write the fully discrete scheme (1.15) in the following *pointwise form*

$$\overline{\partial} U^n(\boldsymbol{x}) + A^n U^n(\boldsymbol{x}) = P_0^n \tilde{f}^n(\boldsymbol{x}),\ \forall \boldsymbol{x} \in \Omega. \tag{2.19}$$

Indeed, in view of $\overline{\partial} U^n + A^n U^n - P_0^n \tilde{f}^n \in \mathbb{V}^n$, (1.15), and (2.5), we have

$$\begin{aligned}\left\langle A^n U^n + \overline{\partial} U^n - P_0^n \tilde{f}^n, \phi \right\rangle &= \left\langle A^n U^n + \overline{\partial} U^n - P_0^n \tilde{f}^n, P_0^n \phi \right\rangle \\ &= a\left(U^n, P_0^n \phi\right) + \left\langle \tau_n^{-1}(U^n - U^{n-1}) - f^n, P_0^n \phi \right\rangle = 0,\end{aligned} \tag{2.20}$$

for any $\phi \in \mathrm{H}_0^1(\Omega)$. Therefore the function $\overline{\partial} U^n + A^n U^n - P_0^n \tilde{f}^n$ vanishes.

2.6. **Error equation.** Let us consider the *(full) error*, the *elliptic reconstruction error* and the *parabolic error* which are defined, respectively as follows

$$e = U - u, \tag{2.21}$$

$$\epsilon = \omega - U, \tag{2.22}$$

$$\rho = \omega - u. \tag{2.23}$$

We have the following decomposition of the error

$$e = \rho - \epsilon. \tag{2.24}$$

We can also readily derive the following error relation for the parabolic error in terms of the reconstruction error and the reconstruction itself [LM06]:

$$\begin{aligned}\left\langle \partial_t \rho(t), \phi \right\rangle + a\left(\rho(t), \phi\right) &= \left\langle \partial_t \epsilon(t), \phi \right\rangle + a\left(\omega(t) - \omega^n, \phi\right) \\ &\quad + \tau_n^{-1} \left\langle P_0^n U^{n-1} - U^{n-1}, \phi \right\rangle + \left\langle P_0^n \tilde{f}^n - f(t), \phi \right\rangle\end{aligned} \tag{2.25}$$

for all $\phi \in \mathrm{H}_0^1(\Omega)$, $t \in I_n$ and $n \in [1:N]$.

3. The duality approach to a posteriori error estimates

In this section we expose the combination of the elliptic reconstruction and the duality techniques in the simple situation where the scheme is discrete only in space. This should help understand the main ideas that we will employ for the fully discrete case in §4.

3.1. **Spatially semi-discrete scheme and continuous time notation.** In this section, since we will be dealing with the space semidiscrete scheme only, we will use the same symbols introduced for the fully discrete scheme in §2, albeit in their semidiscrete analog by dropping the index n. *The notation now introduced is valid only in 3* and it is important not to confuse time-dependent functions, such as U, ω, e, ϵ and ρ, to be introduced next with their fully-discrete analogs introduced in 2. Specifically, let $\mathbb{V}$ be a given (time-invariant) finite element space, we define $U : [0,T] \to \tilde{\mathbb{V}}$ to be the solution of the *semi-discrete Galerkin finite element equation*

$$\begin{gathered} U(0) := Iu(0), \text{ and} \\ \left\langle \partial_t U(t), \Phi \right\rangle + a\left(U(t), \Phi\right) = \left\langle f(t), \Phi \right\rangle,\ \forall \Phi \in \mathbb{V}, t \in [0,T], \end{gathered} \tag{3.1}$$

where the operator I is a suitable interpolation or projection operator from $\mathrm{H}^1(\Omega)$, or $\mathrm{L}_2(\Omega)$, onto $\mathbb{V}$. Following on this definition we define the error at time t to be $e(t) := U(t) - u(t)$ and the semi-discrete elliptic reconstruction to be $\omega(t) := \mathscr{R}U(t)$, were $\mathscr{R}$ is the elliptic reconstruction operator associated with the space $\mathbb{V}$ (compare with 2.2). In analogy with the fully-discrete notation in 2.4, we define the elliptic reconstruction error $\epsilon := \omega - U$ and the parabolic error $\rho := U - u$.

We observe that while in the semi-discrete setting one assumes the discrete space, $\mathbb{V}$, to be invariant in time, in the fully discrete setting, analyzed later in 4, we will take into account the possibility of the discrete space to change, with respect to the timestep. For instance, in an adaptive mesh refinement scheme the space change derives from the mesh's modification from a time to the next.

The notation here introduced leads to the semi-discrete analog of the error equation (2.25)

$$\langle \partial_t \rho(t), \phi \rangle + a(\rho(t), \phi) = \langle \partial_t \epsilon(t), \phi \rangle + \langle P_0 f(t) - f(t), \phi \rangle, \tag{3.2}$$

for $\phi \in \mathrm{H}^1_0(\Omega), t \in (0, T]$. This equation consitutes the starting point for the estimate derivation.

3.2. **The dual solution.** The concept of dual solution, with that of elliptic reconstruction, consitute the two main analytical tool of this paper. For each $s \leq T$, we define the *dual solution* to be the function

$$z(x, t; s) = z_s(x, t), \text{ for } x \in \Omega \text{ and } 0 \leq t \leq s, \tag{3.3}$$

which satisfies $z_s \in \mathrm{L}_2(0, T; \mathrm{H}^1_0(\Omega))$, $\partial_t z_s \in \mathrm{L}_2(0, T; \mathrm{H}^{-1}(\Omega))$, and solves the following backward parabolic *dual problem*:

$$\begin{aligned} -\langle \partial_t z_s(t), \phi \rangle + a(\phi, z_s(t)) = 0, \ \forall \phi \in \mathrm{H}^1_0(\Omega), t \in [0, s), \\ z_s(x, s) = \rho(x, s), \ \forall x \in \Omega \end{aligned} \tag{3.4}$$

for each $s \in [0, T]$. Notice that ϕ can be taken to be time dependent, with the appropriate differentiability properties.

The dual solution enjoys stability properties which we will use in the sequel. An immediate property is the usual energy identity

$$\|z_s(t)\|^2 + 2 \int_t^s |z_s|_a^2 = \|\rho(s)\|^2, \ \forall t \in [0, s]. \tag{3.5}$$

A more intricate property is given by the following result.

3.3. **Lemma** (Strong stability estimate [EJ91, Lem. 4.2]) *Suppose that Ω is a convex domain. Then for each $s \in [0, T]$ it holds*

$$\max \left(\int_0^s \|\partial_t z_s(t)\|^2 (s - t)\, \mathrm{d}t, \int_0^s \|-\operatorname{div}(\boldsymbol{A} \nabla z_s)(t)\|^2 (s - t)\, \mathrm{d}t \right) \leq \frac{1}{4} \|\rho(s)\|^2. \tag{3.6}$$

3.4. **Error analysis in the spatially semi-discrete case.** Integrating by parts in time, problem (3.4) implies

$$\begin{aligned} \langle \rho(s), \phi(s) \rangle &= \langle z_s(s), \phi(s) \rangle \\ &= \langle z_s(0), \phi(0) \rangle + \int_0^s \langle \partial_t \phi(t), z_s(t) \rangle + a(\phi(t), z_s(t))\, \mathrm{d}t \end{aligned} \tag{3.7}$$

for all $\phi \in \mathrm{L}_2(0, T; \mathrm{H}^1_0(\Omega))$ such that $\partial_t \phi \in \mathrm{L}_2(0, T; \mathrm{H}^{-1}(\Omega))$.

Taking $\phi = \rho$, using (3.2) and assuming $P_0 \tilde{f} - f = 0$ momentarily—in the proof of Theorem 4.2 we remove this assumption—we obtain

$$\|\rho(s)\|^2 = \langle \rho(0), z_s(0) \rangle + \int_0^s \langle \partial_t \epsilon(t), z_s(t) \rangle\, \mathrm{d}t. \tag{3.8}$$

The first term on the right-hand side, is easily estimated, with Lemma 3.3 in mind, as follows

$$\langle \rho(0), z_s(0)\rangle \le \|\rho(0)\| \sup_{[0,s]} \|z_s\| . \tag{3.9}$$

As for the second term on the right-hand side of (3.8) we have the choice of two different ways for estimating it.

(a) A direct estimate yields

$$\int_0^s \langle \partial_t \epsilon, z_s\rangle \le \sup_{[0,s]} \|z_s\| \int_0^s \|\partial_t \epsilon\| . \tag{3.10}$$

Notice that the term $\partial_t \epsilon$ can be estimated via elliptic a posteriori error estimates because it is the difference between $\partial_t U$ and its reconstruction $\mathscr{R}\partial_t U = \partial_t \mathscr{R} U$.

(b) A less direct estimate, that would avoid the appearance of time derivatives in the estimator, is obtained by integrating by parts in time first

$$\int_0^s \langle \partial_t \epsilon, z_s\rangle = \langle \epsilon(s), z_s(s)\rangle - \langle \epsilon(0), z_s(0)\rangle - \int_0^s \langle \epsilon(t), \partial_t z_s(t)\rangle \,\mathrm{d}t. \tag{3.11}$$

The last integral can be then bounded as follows

$$\begin{aligned} \int_0^s \epsilon(t)\partial_t z_s(t)\,\mathrm{d}t &\le \int_0^s \frac{\|\epsilon(t)\|}{\sqrt{s-t}} \|\partial_t z_s(t)\| \sqrt{s-t}\,\mathrm{d}t \\ &\le \left(\int_0^s \frac{\|\epsilon(t)\|^2}{s-t}\,\mathrm{d}t\right)^{1/2} \left(\int_0^s \|\partial_t z_s(t)\|^2 (s-t)\,\mathrm{d}t\right)^{1/2} . \end{aligned} \tag{3.12}$$

Unfortunately this bound may not be useful. Indeed, for the first integral in the right-hand side to be finite it is necessary that $\epsilon(t) = \mathrm{o}(1)$ at $t = s$. This means that the error between the discrete solution and its reconstruction should at least vanish at s. Heuristically this can be interpreted as the mesh having to become infinitely fine as time gets closer to s: an unrealistic option.

To circumvent this difficulty, without totally sacrificing $\|\epsilon\|$ to $\|\partial_t \epsilon\|$, we compromise between approach (a) and (b) by following through from (3.8) as follows: fix $r \in (0, s)$ (think of it as a close point to s), split the integral and integrate by parts in time

$$\begin{aligned} \|\rho(s)\|^2 &= \langle z_s(0), \rho(0)\rangle + \left(\int_0^r + \int_r^s\right) \langle \partial_t \epsilon, z_s\rangle \\ &= \langle z_s(0), \rho(0) - \epsilon(0)\rangle + \langle z_s(r), \epsilon(r)\rangle - \int_0^r \langle \epsilon, \partial_t z_s\rangle + \int_r^s \langle \partial_t \epsilon, z_s\rangle \\ &\le \sup_{[0,s]} \|z_s\| \left(\|e(0)\| + \|\epsilon(r)\| + \int_r^s \|\partial_t \epsilon\|\right) \\ &\quad + \left(\int_0^r \|\partial_t z_s(t)\|^2 (s-t)\,\mathrm{d}t\right)^{1/2} \left(\int_0^r \frac{\|\epsilon(t)\|^2}{s-t}\,\mathrm{d}t\right)^{1/2} . \end{aligned} \tag{3.13}$$

The stability estimates (3.5) and (3.6) imply that

$$\|\rho(s)\| \le \|e(0)\| + \|\epsilon(r)\| + \int_r^s \|\partial_t \epsilon\| + \frac{1}{2}\left(\int_0^r \frac{\|\epsilon(t)\|^2}{s-t}\,\mathrm{d}t\right)^{1/2} . \tag{3.14}$$

We have essentially proved the following result.

3.5. **Theorem** (Semi-discrete duality-reconstruction a posteriori error estimate) *Suppose that* $f(t) \in \tilde{\mathbb{V}}$*, for* $t \in [0,T]$*, and that there exists an a posteriori elliptic error estimator function* $\mathscr{E}[\cdot,\cdot,\cdot]$ *such that*

$$\|v - \mathscr{R}v\| \le \mathscr{E}[v, \tilde{\mathbb{V}}, \mathrm{L}_2(\Omega)], \text{ for } v \in \mathrm{H}_0^1(\Omega), \tag{3.15}$$

then the error occuring in the semi-discrete scheme (3.1) *obeys the a posteriori bound*

$$\begin{aligned}\sup_{t\in[0,s]} \|U(t) - u(t)\| \le \|U(0) - u(0)\| + L(s,r) \sup_{[0,r]} \mathscr{E}[U, \tilde{\mathbb{V}}, \mathrm{L}_2(\Omega)] \\ + (s-r) \sup_{[r,s]} \mathscr{E}[\partial_t U, \tilde{\mathbb{V}}, \mathrm{L}_2(\Omega)]\end{aligned} \tag{3.16}$$

where

$$L(s,r) := \frac{1}{2}\sqrt{\log \frac{s}{s-r}}. \tag{3.17}$$

Proof Fix r and s, use (3.14) with $s' < s$ instead of s, $r' = s' - (s-r)$ instead of r. Take max for s'. ∎

3.6. **Corollary** (Semi-discrete duality-residual a posteriori estimates) *If* Ω *is a convex domain in* $\mathbb{R}^d$ *and* $f(t) \in \mathrm{L}_2(\Omega)$ *for each* $t \in [0,T]$*, then the following a posteriori error estimate holds*

$$\begin{aligned}\sup_{[0,s]} \|U - u\| \le \|U(0) - u(0)\| \\ + L(s,r) \sup_{[0,r]} \Big(C_3 \left\|h^2(AU - A_h U)\right\| \\ + C_5 \left\|h^{3/2} J[U]\right\|_\Sigma + C_7 \left\|h^2(P_0\tilde{f} - f)\right\| \Big) \\ + (s-r) \sup_{[r,s]} \Big(C_3 \left\|h^2(A\partial_t U - A_h \partial_t U)\right\| \\ + C_5 \left\|h^{3/2} J[\partial_t U]\right\|_\Sigma + \frac{1}{2\sqrt{\alpha}} \left\|h(P_0\tilde{f} - f)\right\| \Big).\end{aligned} \tag{3.18}$$

Proof Sketch: apply the Theorem and add the f part. ∎

4. Estimates for the fully discrete scheme

Having understood the main idea of the techinque in the last section, we turn our attention to the analysis of the fully discrete scheme (1.15). For convenience, we switch notation slightly and use the symbol U, even without the superscript n, for the fully discrete solution and its piecewice linear interpolation now.

We introduce first some extra notation to be used in this section.

4.1. **Definition** (Error indicators and estimators) Suppose an a posteriori elliptic error estimator function $\mathscr{E}[\cdot,\cdot,\cdot]$ is available, introduce the following (time-local) *$\mathscr{E}$-based spatial error estimators*

$$\varepsilon_n = \mathscr{E}[U^n, \tilde{\mathbb{V}}^n, \mathrm{L}_2(\Omega)], \tag{4.1}$$

$$\eta_n = \mathscr{E}[\partial_t U^n, \tilde{\mathbb{V}}^n \cap \tilde{\mathbb{V}}^{n-1}, \mathrm{L}_2(\Omega)], \tag{4.2}$$

and the *time estimator*

$$\theta_n = \left\|A^{n-1}U^{n-1} - A^n U^n\right\| = \begin{cases} \frac{1}{2}\left\|P_0^1 \tilde{f}^1 - \overline{\partial} U^1 - A^0 U^0\right\| & \text{for } n = 1, \\ \frac{1}{2}\left\|\partial\left(P_0^n \tilde{f}^n - \overline{\partial} U^n\right)\right\| \tau_n & \text{for } n \in [2:N], \end{cases} \tag{4.3}$$

and the *logarithmic time accumulation coefficients*

$$b_n = \begin{cases} \frac{1}{4}\log\left(\frac{T-t_{n-1}}{T-t_n}\right) & \text{for } n \in [1:N-1], \\ \frac{1}{8} & \text{for } n = N; \end{cases} \tag{4.4}$$

$$a_n = \begin{cases} \lambda\left(\frac{\tau_n}{T-t_n}\right) - \lambda\left(-\frac{\tau_{n+1}}{T-t_n}\right), & \text{for } n \in [0:n-2], \\ \lambda\left(\frac{\tau_{N-1}}{\tau_N}\right) - 1, & \text{for } n = N-1, \end{cases} \tag{4.5}$$

where

$$\lambda(x) := \begin{cases} (1+1/x)\log(1+x) & \text{for } |x| \in (0,1), \\ 1 & \text{for } x = 0, \end{cases} \tag{4.6}$$

which is an increasing function of x. We observe that the functions $\lambda(x) - 1$, $1 - \lambda(-x)$ and $\lambda(x) - \lambda(-y)$ are positive for $(x, y) \in (0,1)^2$, a fact that makes the coefficients a_n to be positive.

Finally we introduce the *data approximation* estimators

$$\beta_n = \int_{t_{n-1}}^{t_n} \|f^n - f(t)\| \,\mathrm{d}t; \tag{4.7}$$

and we notice that for f regular enough we can replace β_n by the right hand side of the inequality

$$\int_{t_{n-1}}^{t_n} \|f^n - f(t)\| \,\mathrm{d}t \le \|\partial_t f\|_{\mathrm{L}_1(I_n,\mathrm{L}_2(\Omega))} \tau_n. \tag{4.8}$$

4.2. **Theorem** (General duality a posteriori parabolic error estimate) *Let u be the exact solution of (1.12), $(U^n)_{n\in[1:N]}$ the corresponding (fully) discrete solution of (1.15) and $(\omega^n) = (\mathscr{R}^n U)$ the elliptic reconstruction of U, as defined by (2.8), then the following a posteriori error estimate holds*

(4.9)

$$\begin{aligned} \left\|\omega^N - u(T)\right\| \le \|U(0) - u(0)\| &+ \left(\sum_{n=0}^{N-1} a_n \varepsilon_n^2\right)^{1/2} + \eta_N \\ &+ \left(\sum_{n=1}^{N} b_n \left(\left\|U^{n-1} - U^n\right\| + \eta_n\right)^2\right)^{1/2} + \sum_{n=1}^{N} \tau_n \|\partial_t f\|_{\mathrm{L}_1(I_n;\mathrm{L}_2(\Omega))}, \\ &+ \sqrt{\frac{\tau_N}{2}} \|\gamma_N h_N\| + \left(\sum_{n=1}^{N-1} b_n \left\|\gamma_n h_n^2\right\|^2\right)^{1/2} \end{aligned}$$

where the various error estimators and the coefficients ε_n, η_n, a_n, b_n, β_n, and γ_n are defined in §4.1.

An alternative time error estimator is given by replacing the fourth summand on the right hand side by (4.9) by

$$\left(\sum_{n=1}^{N} b_n \theta n^2\right)^{1/2}. \tag{4.10}$$

4.3. **Corollary** (Duality a posteriori full error estimates) *With the same notation as in 4.2 we have*

$$
\begin{aligned}
\left\|U^N - u(T)\right\| \le \left\|U^0 - u(0)\right\| + \sqrt{\frac{\tau_N}{2}}\left\|\gamma_N h_N\right\| + \sum_{n=1}^{N} \tau_n \left\|\partial_t f\right\|_{\mathrm{L}_1(I_n;\mathrm{L}_2(\Omega))} \\
+ \sqrt{1+\log\frac{T}{\tau_N}}\Bigg(\max_{n\in[0:N]} \varepsilon_n + 2 \max_{n\in[1:N-1]} \left\|\gamma_n h_n^2\right\| \\
+ \frac{1}{2}\max_{n\in[1:N]}\left(\left\|U^{n-1}-U^n\right\| + \eta_n\right)\Bigg).
\end{aligned}
\tag{4.11}
$$

Furthermore, we can eliminate all the terms ε_n from the estimate by replacing the last term by

$$
\max_{n\in[1:N]} \theta_n. \tag{4.12}
$$

4.4. **Remark** (comparison between Theorem 4.2 and Corollary 4.3) Corollary 4.3 has a simpler estimate than Theorem 4.2 in that it involves less terms and does not require as much memory. Notice however, that from an error bound view-point, the Theorem's tighter bound may be more effective as the time accumulation is not as strict as in the Corollary. This is especially true in problems, typical in the parabolic setting, where the initial error may be very big and gets damped with time.

5. Proof of the main results

As with the semi-discrete case of §3, the starting point is the fully discrete analog of (3.8), which is readily obtained from (2.25) and (3.4), and reads

$$
\begin{aligned}
\left\|\rho(T)\right\|^2 = \langle \rho(0), z_T(0)\rangle + \int_0^T \langle \partial_t \epsilon(t), z_T(t)\rangle \,\mathrm{d}t \\
+ \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} a\left(\omega(t)-\omega^n, z_T(t)\right) + \langle \gamma_n, z_T(t)\rangle + \left\langle \tilde f^n - f(t), z_T(t)\right\rangle \,\mathrm{d}t.
\end{aligned}
\tag{5.1}
$$

where

$$
\gamma_n := \frac{P_0^n U^{n-1} - U^{n-1}}{\tau_n} + P_0^n \tilde f^n - \tilde f^n = (P_0^n - \mathrm{I})(\tau_n^{-1} U^{n-1} + \tilde f^n). \tag{5.2}
$$

5.1. **Space error estimate.** The first two terms are estimated, similarly to (3.13), as follows

$$
\begin{aligned}
\langle \rho(0), z_T(0)\rangle + \int_0^T \langle \partial_t \epsilon(t), z_T(t)\rangle \,\mathrm{d}t \le \\
\left\|\rho(T)\right\| \left(\left\|e(0)\right\| + \left\|\epsilon(t_{N-1})\right\| + \int_{t_{N-1}}^T \left\|\partial_t \epsilon\right\| + \frac{1}{2}\left(\int_0^{t_{N-1}} \frac{\left\|\epsilon(t)\right\|^2}{T-t}\,\mathrm{d}t\right)^{1/2}\right).
\end{aligned}
\tag{5.3}
$$

To proceed we observe that

$$
\int_{t_{N-1}}^T \left\|\partial_t \epsilon\right\| = \left\|\epsilon^N - \epsilon^{N-1}\right\| \le \eta_N. \tag{5.4}
$$

and that, by convexity and linearity,

$$
\begin{aligned}
\int_0^{t_{N-1}} \frac{\|\epsilon(t)\|^2}{T-t}\,\mathrm{d}t &= \int_0^{t_{N-1}} \frac{\left\|\sum_{n=0}^{N-1} \epsilon^n l_n(t)\right\|^2}{T-t}\,\mathrm{d}t \\
&\leq \sum_{n=0}^{N-1} \|\epsilon^n\|^2 \int_0^{t_{N-1}} \frac{l_n(t)}{T-t}\,\mathrm{d}t = \sum_{n=0}^{N-1} a_n \varepsilon_n^2.
\end{aligned} \tag{5.5}
$$

Thus we obtain

$$
\langle \rho(0), z_T(0)\rangle + \int_0^T \langle \partial_t \epsilon(t), z_T(t)\rangle\,\mathrm{d}t \leq \|\rho(T)\| \left(\|e(0)\| + \left(\sum_{n=0}^{N-1} a_n \varepsilon_n^2 \right)^{1/2} + \eta_N \right). \tag{5.6}
$$

We notice that the former estimate implies the more traditional one [EJ91]

$$
\begin{aligned}
\int_0^{t_{N-1}} \frac{\|\epsilon(t)\|^2}{T-t}\,\mathrm{d}t &\leq \max_{n\in[0:N-1]} \|\epsilon^n\|^2 \left(1 + \sum_{n=1}^{N-1} a_n\right) \\
&= \left(1 + 4L(T, t_{N-1})^2\right) \max_{n\in[0:N-1]} \|\epsilon^n\|^2
\end{aligned} \tag{5.7}
$$

where $L(T, t_{N-1})$ is the logarithmic factor defined in (3.17).

5.2. **Time error estimate.** The third term in (5.1), which accounts mainly for the time error, can be bounded as follows

$$
\begin{aligned}
\sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} a\left(\omega(t) - \omega^n, z_T(t)\right)\,\mathrm{d}t &\leq \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \left\|\omega^{n-1} - \omega^n\right\| l_{n-1}(t) \left\|-\operatorname{div}\left(\boldsymbol{A}\nabla z_T\right)(t)\right\|\,\mathrm{d}t \\
&\leq \frac{1}{2}\|\rho(T)\| \left(\sum_{n=1}^{N} \left\|\omega^{n-1} - \omega^n\right\|^2 \int_{t_{n-1}}^{t_n} \frac{l_{n-1}(t)^2}{T-t}\,\mathrm{d}t \right)^{1/2} \\
&\leq \frac{1}{2}\|\rho(T)\| \left(\frac{1}{2}\left\|\omega^N - \omega^{N-1}\right\|^2 + \sum_{n=1}^{N-1} \left\|\omega^n - \omega^{n-1}\right\|^2 \log\left(\frac{T - t_{n-1}}{T - t_n}\right) \right)^{1/2}.
\end{aligned} \tag{5.8}
$$

Notice that we do not need to deal with the last interval $(t_{N-1}, T]$ as a special case since $l_{N-1}(t) = \mathrm{O}(T-t)$ compensates for the singularity of $1/(T-t)$. The terms $\left\|\omega^{n-1} - \omega^n\right\|$ appearing in this estimate still need to be estimated, as there is no explicit knowledge of the reconstructed functions $\omega^n = \mathscr{R}^n U^n$. These terms can be dealt with in two different ways.

(a) One way to estimate these terms is given by:

$$
\begin{aligned}
\left\|\omega^{n-1} - \omega^n\right\| &\leq \left\|U^{n-1} - U^n\right\| + \left\|\omega^{n-1} - \omega^n - U^{n-1} + U^n\right\| \\
&= \left\|U^{n-1} - U^n\right\| + \tau_n \left\|\partial_t \epsilon^n\right\| \leq \left\|U^{n-1} - U^n\right\| + \eta_n,
\end{aligned} \tag{5.9}
$$

for all $t \in I_n$. Thus we obtain the estimate

$$
\sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} a\left(\omega(t) - \omega^n, z_T(t)\right)\,\mathrm{d}t \leq \|\rho(T)\| \left(\sum_{n=1}^{N} b_n \left(\left\|U^{n-1} - U^n\right\| + \eta_n\right)^2 \right)^{1/2}. \tag{5.10}
$$

(b) Another way to estimate $\left\|\omega^{n-1}-\omega^n\right\|$ consists in using again the definition of elliptic reconstruction and the Poincaré inequality as follows:

$$
\begin{aligned}
\left\|\omega^{n-1}-\omega^n\right\|^2 &\leq C_{2,1} a\left(\omega^{n-1}-\omega^n, \omega^{n-1}-\omega^n\right) \\
&= C_{2,1}\left\langle A^{n-1}U^{n-1}-A^nU^n, \omega^{n-1}-\omega^n\right\rangle \\
&\leq C_{2,1}\left\|A^{n-1}U^{n-1}-A^nU^n\right\|\left\|\omega^{n-1}-\omega^n\right\|,
\end{aligned}
\tag{5.11}
$$

thus obtaining

$$
\left\|\omega^{n-1}-\omega^n\right\| \leq C_{2,1}\left\|A^{n-1}U^{n-1}-A^nU^n\right\| = \theta_n. \tag{5.12}
$$

Therefore

$$
\sum_{n=1}^N \int_{t_{n-1}}^{t_n} a\left(\omega(t)-\omega^n, z_T(t)\right) \,\mathrm{d}t \leq \|\rho(T)\|\left(\sum_{n=1}^N b_n\theta_n^2\right)^{1/2}. \tag{5.13}
$$

Also here, the estimator simplifies if we relax the bound by taking the maximum norm in time

$$
\sum_{n=1}^N \int_{t_{n-1}}^{t_n} a\left(\omega(t)-\omega^n, z_T(t)\right) \,\mathrm{d}t \leq \|\rho(T)\| \max_{n\in[1:N]}\left\|\omega^n-\omega^{n-1}\right\|\sqrt{\left(\frac{1}{8}+L(T,t_{N-1})^2\right)}. \tag{5.14}
$$

5.3. **Remark** The above bound could be derived differently by using not the strong stability, but rather the definition of elliptic reconstruction directly. We refer to our companion paper for a detailed estimate of this type [LM06, § 3.5].

5.4. **Data approximation and coarsening error estimates.** We bound now the third term in (5.1):

$$
\begin{aligned}
\sum_{n=1}^N \int_{t_{n-1}}^{t_n} \left\langle\gamma_n, z_T(t)\right\rangle \,\mathrm{d}t &= \sum_{n=1}^N \int_{t_{n-1}}^{t_n} \left\langle\gamma_n, z_T(t)-\Pi^n z_T(t)\right\rangle \,\mathrm{d}t \\
&\leq \sum_{n=1}^{N-1} \int_{t_{n-1}}^{t_n} \left\|\gamma_n h_n^2\right\| \left|z_T(t)\right|_2 \,\mathrm{d}t + \int_{t_{N-1}}^{T} \left\|\gamma_N h_N\right\| \left|z_T(t)\right|_1 \,\mathrm{d}t \\
&\leq \frac{1}{2}\|\rho(T)\|\left(\sqrt{2\tau_N}\left\|\gamma_N h_N\right\| + \left(\sum_{n=1}^{N-1}\left\|\gamma_n h_n^2\right\|^2 \log\left(\frac{T-t_{n-1}}{T-t_n}\right)\right)^{1/2}\right) \\
&= \|\rho(T)\|\left(\sqrt{\frac{\tau_N}{2}}\left\|\gamma_N h_N\right\| + \left(\sum_{n=1}^{N-1} b_n\left\|\gamma_n h_n^2\right\|^2\right)^{1/2}\right).
\end{aligned}
\tag{5.15}
$$

Here we have used the fact that Ω is convex in order to apply the estimate

$$
\left|z_T(t)\right|_2 \leq \left\|-\operatorname{div}\left(\boldsymbol{A}\nabla z_T\right)(t)\right\|, \tag{5.16}
$$

and then apply the strong stability estimate (3.6). As with the space and time estimates, this estimate can be simplified, with some loss of sharpness, as follows

$$
\begin{aligned}
&\sum_{n=1}^N \int_{t_{n-1}}^{t_n} \left\langle\gamma_n, z_T(t)\right\rangle \,\mathrm{d}t \\
&\qquad \leq \|\rho(T)\|\left(\sqrt{\frac{\tau_N}{2}}\left\|\gamma_N h_N\right\| + L(T,t_{N-1}) \max_{n\in[1:N-1]}\left\|\gamma_n h_n^2\right\|\right).
\end{aligned}
\tag{5.17}
$$

The fourth term in (5.1) can be bounded in two different ways depending on which definition for $\tilde{f}^n$ is taken.

(a) If $\tilde{f}^n = f^n$ then we can proceed as follows

$$\begin{aligned}\sum_{n=1}^{N}\int_{t_{n-1}}^{t_n}\left\langle \tilde{f}^n - f(t), z_T(t)\right\rangle \mathrm{d}t &\le \sum_{n=1}^{N}\max_{I_n}\|z_T\|\int_{t_{n-1}}^{t_n}\left\|\tilde{f}^n - f(t)\right\| \mathrm{d}t \\ &\le \|\rho(T)\|\sum_{n=1}^{N}\beta_n.\end{aligned} \tag{5.18}$$

(b) If instead we have $\tilde{f}^n = \int_{t_{n-1}}^{t_n} f(t)\,\mathrm{d}t/\tau_n$ then we can exploit the orthogonality in time and write, for each $n \in [1 : N-1]$

$$\begin{aligned}\int_{t_{n-1}}^{t_n}\left\langle \tilde{f}^n - f(t), z_T(t)\right\rangle \mathrm{d}t &= \int_{t_{n-1}}^{t_n}\left\langle \tilde{f}^n - f(t), z_T(t) - z_T(t_{n-1})\right\rangle \mathrm{d}t \\ &\le \max_{t\in I_n}\|z_T(t) - z_T(t_{n-1})\|\int_{t_{n-1}}^{t_n}\left\|\tilde{f}^n - f(t)\right\| \mathrm{d}t.\end{aligned} \tag{5.19}$$

By noticing that

$$\begin{aligned}\max_{t\in I_n}\|z_T(t) - z_T(t_{n-1})\| &= \max_{t\in I_n}\left\|\int_{t_{n-1}}^{t}\partial_s z_T(s)\,\mathrm{d}s\right\| \\ &\le \max_{t\in I_n}\int_{t_{n-1}}^{t}\|\partial_s z_T(s)\|\,\mathrm{d}s \le \int_{t_{n-1}}^{t_n}\|\partial_t z_T\| \\ &\le \log\left(\frac{T - t_{n-1}}{T - t_n}\right)^{1/2}\left(\int_{t_{n-1}}^{t_n}\|\partial_t z_T(t)\|^2 (T-t)\,\mathrm{d}t\right)^{1/2} \\ &= 2b_n^{1/2}\left(\int_{t_{n-1}}^{t_n}\|\partial_t z_T(t)\|^2 (T-t)\,\mathrm{d}t\right)^{1/2}.\end{aligned} \tag{5.20}$$

Summing up and using the strong stability estimate (3.6) we obtain

$$\sum_{n=1}^{N}\int_{t_{n-1}}^{t_n}\left\langle \tilde{f}^n - f(t), z_T(t)\right\rangle \mathrm{d}t \le \|\rho(T)\|\left(\beta_N + 2\left(\sum_{n=1}^{N-1} b_n\beta_n^2\right)^{1/2}\right). \tag{5.21}$$

Like earlier estimates, this estimate can be further simplified, by taking the maximum and slightly relaxing it, into

$$\begin{aligned}&\sum_{n=1}^{N}\int_{t_{n-1}}^{t_n}\left\langle \tilde{f}^n - f(t), z_T(t)\right\rangle \mathrm{d}t \\ &\qquad\le \|\rho(T)\|\left(\beta_N + 2L(T, t_{N-1})\max_{n\in[1:N-1]}\beta_n\right).\end{aligned} \tag{5.22}$$

Replacing the estimates (5.6), (5.10)/ (5.13) , (5.15) and (5.18) into the relation (5.1) we obtain the result of Theorem 4.2.

5.5. **Proof of Corollary 4.3.** The splitting $\rho = \epsilon + e$ and some simple Hölder inequality manipulations on the summations on the logarithmic coefficients yield the result.

6. Sample application: residual estimates in space

6.1. **Definition** (residuals) The residuals constitute the building blocks of the a posteriori estimators used in this paper. We associate with equations (1.12) and (2.19) two residual functions: the *inner residual* is defined as

$$\begin{aligned} R^0 &:= A_{\mathrm{el}} U^0 - A^0 U^0, \\ R^n &:= A_{\mathrm{el}} U^n - A^n U^n = A_{\mathrm{el}} U^n - P_0^n \tilde{f}^n + \overline{\partial} U^n, \text{ for } n \in [1:N], \end{aligned} \tag{6.1}$$

and the *jump residual* which is defined as

$$J^n := J[U^n] = [\![\nabla U^n]\!] . \tag{6.2}$$

With definition §2.1 in mind, the inner residual terms can be written explicitly as

$$\langle R^n, \phi \rangle = \sum_{K \in \mathscr{T}_n} \left\langle -\operatorname{div}(\boldsymbol{A} \nabla v) - P_0^n \tilde{f}^n + \frac{U^n - P_0^n U^{n-1}}{\tau_n}, \phi \right\rangle_K . \tag{6.3}$$

We can now introduce, for $n \in [0:N]$, the *elliptic reconstruction error estimators*

$$\varepsilon_n := C_{6,2} \left\| h_n^2 R^n \right\| + C_{10,2} \left\| h_n^{3/2} J^n \right\|_{\Sigma_n} , \tag{6.4}$$

and, for $n \in [1:N]$, the *space error estimator*

$$\eta_n := C_{6,2} \left\| \hat{h}_n^2 \partial R^n \right\| + C_{10,2} \left\| \hat{h}_n^{3/2} \partial J^n \right\|_{\hat{\Sigma}_n} + C_{14,2} \left\| \hat{h}_n^{3/2} \partial J^n \right\|_{\check{\Sigma}_n \setminus \hat{\Sigma}_n} . \tag{6.5}$$

6.2. **Conclusion.** With these definitions, Corollary 4.3 includes the estimates of Eriksson & Johnson [EJ91], but with no unrealistic assumptions on the meshes from time to time.

Note also that a direct application of Theorem 4.2 provides finer estimates with respect to time accumulation; these are especially helpful in situations where the error (on a time-invariant mesh) decreases with time.

References

[AMN05] Georgos Akrivis, Charalambos Makridakis, and Ricardo H. Nochetto, *A posteriori error estimates for the Crank–Nicolson method for parabolic equations*, Math. Comp (to appear 2005).

[AO00] Mark Ainsworth and J. Tinsley Oden, *A posteriori error estimation in finite element analysis*, Wiley-Interscience [John Wiley & Sons], New York, 2000. MR 1 885 308

[BBM05] A. Bergam, C. Bernardi, and Z. Mghazli, *A posteriori analysis of the finite element discretization of some parabolic equations*, Math. Comp. (2005).

[Bra01] Dietrich Braess, *Finite elements*, second ed., Cambridge University Press, Cambridge, 2001, Theory, fast solvers, and applications in solid mechanics, Translated from the 1992 German edition by Larry L. Schumaker. MR 2001k:65002

[BS94] Susanne C. Brenner and L. Ridgway Scott, *The mathematical theory of finite element methods*, Springer-Verlag, New York, 1994. MR 95f:65001

[BV04] Christine Bernardi and Rüdiger Verfürth, *A posteriori error analysis of the fully discretized time-dependent Stokes equations*, M2AN Math. Model. Numer. Anal. **38** (2004), no. 3, 437–455. MR MR2075754 (2005g:65131)

[Cia78] Philippe G. Ciarlet, *The finite element method for elliptic problems*, North-Holland Publishing Co., Amsterdam, 1978, Studies in Mathematics and its Applications, Vol. 4. MR 58 #25001

[EJ91] Kenneth Eriksson and Claes Johnson, *Adaptive finite element methods for parabolic problems. I. A linear model problem*, SIAM J. Numer. Anal. **28** (1991), no. 1, 43–77. MR 91m:65274

[GAT00] Bosco García-Archilla and Edriss S. Titi, *Postprocessing the Galerkin method: the finite-element case*, SIAM J. Numer. Anal. **37** (2000), no. 2, 470–499 (electronic). MR 2001h:65112

[LM06] Omar Lakkis and Charalambos Makridakis, *Elliptic reconstruction and a posteriori error estimates for fully discrete linear parabolic problems*, Math. Comp. **75** (2006), no. 256, 1627–1658 (electronic). MR MR2240628 (2007e:65122)

[LN03] Xiaohai Liao and Ricardo H. Nochetto, *Local a posteriori error estimates and adaptive control of pollution effects*, Numer. Methods Partial Differential Equations **19** (2003), no. 4, 421–442. MR MR1980188 (2004c:65130)

[MN03] Charalambos Makridakis and Ricardo H. Nochetto, *Elliptic reconstruction and a posteriori error estimates for parabolic problems*, SIAM J. Numer. Anal. **41** (2003), no. 4, 1585–1594 (electronic). MR MR2034895 (2004k:65157)

[Tho97] Vidar Thomée, *Galerkin finite element methods for parabolic problems*, Springer Series in Computational Mathematics, vol. 25, Springer-Verlag, Berlin, 1997. MR 98m:65007

[Ver96] Rüdiger Verfürth, *A review of* a posteriori *error estimation and adaptive mesh-refinement techniques*, Wiley-Teubner, Chichester-Stuttgart, 1996.

Omar Lakkis
Department of Mathematics
University of Sussex
Brighton
UK-BN1 9RF, United Kingdom
E-mail address: omar@iacm.forth.gr

Charalambos Makridakis
Department of Applied Mathematics
University of Crete
GR- 71409 Heraklion, Greece
and
Institute for Applied and Computational Mathematics
Foundation for Research and Technology-Hellas
Vasilika Vouton P.O.Box 1527
GR-71110 Heraklion, Greece
E-mail address: makr@tem.uoc.gr